

Collaboration for Big Data Analytics: Investigating the (Troubled) Relationship between Data Science Experts and Functional Managers

Janine Hagen
LMU Munich
hagen@bwl.lmu.de

Thomas Hess
LMU Munich
thess@bwl.lmu.de

Abstract

The utilization of insights from big data analytics (BDA) in business operations has been identified as a major driver to unlock value from big data. This emphasizes the importance of the involvement of functional business managers in BDA projects and draws attention to their collaboration with BDA experts, such as data scientists. Scholars have identified several challenges that explain why the success rates of BDA projects remain low. However, the relationship between managers and data science experts has not yet been examined as a potential reason for failure. By applying a social capital perspective on the relationship between these groups, we employ a multiple case study to investigate possible obstacles. We find that the relationship is largely troubled due to incongruent cognitive interpretations of BDA applications in the business context, and the absence of structural network ties. These findings suggest a previously under-researched reason why BDA projects still frequently fail.

1. Introduction

Big data analytics (BDA) has been acknowledged as an important driver for business value in the digital age as it can improve agility, innovation, and competitive performance [10, 19]. However, companies seem to struggle to leverage the potential of BDA, as multiple surveys report that most BDA projects fail to deliver business value. Gartner, for example, found a success rate of only 15% for BDA projects and puts forward a similar outlook for the future by forecasting that only 20% of analytics insights will deliver business outcomes through 2022 [3, 38].

Scholars have identified various reasons why BDA projects fail. These are usually grounded in the data itself, in the related process, and in the management of BDA [30]. First, *data challenges* relate to the inherent

characteristics of big data. That is, for example, the enormous scale of the data which makes determining, retrieving, processing, integrating, and inferring the data a challenging task. Second, *process challenges* occur while handling the data. That is, for example, the extraction and cleaning of data from a pool of large-scale unstructured data. Third, *management challenges* occur while accessing and governing the data. These include, for example, issues such as data privacy, security, and a lack of skills to work with the data [30].

To overcome those challenges and to successfully exploit the potential of BDA, different organizational resources are required [18]. These can be divided into *tangible resources* (e.g. technology), *intangible resources* (e.g. data-driven culture), and *human resources* [15]. Human resources comprise technical skills, i.e. the know-how to use new forms of technology to extract intelligence from big data, and managerial skills, i.e. the understanding of how and where to apply the BDA insights in business [15].

Prior research has mainly focused on the tangible resources, e.g. BDA infrastructure and tools, whereas other related aspects, i.e. human resources, have been “largely disregarded” [18, p. 548]. However, human skills and knowledge are highly relevant, as only the combination of data science skills and managerial skills can solve the aforementioned challenges. This argument is in line with IT capability research, where technical and managerial skills have been identified as the critical dimensions of human resources [5, 7]. Additionally, a good working relationship is required between data science experts (e.g. data scientists and data engineers) and other functional managers (e.g. from the marketing or supply chain department) [18]. A fruitful collaboration between these two groups in fact is crucial for the success of BDA projects, as the utilization of data insights in business operations has been identified to be the most critical contributor to unlock the business value of BDA [10].

Considering the unsatisfactory contribution of BDA insights to business outcomes as mentioned in the beginning, we conclude that the required relationship

between data science experts and functional managers may be troubled and needs further investigation. Studies focusing on the management challenges of BDA [30, 35] point out that organizations need to have “the right people, with the right skills” [35, p. 631] in the organization, but say little about the connection of these people. Other studies reveal the skills and personal attributes a data scientist must have, but say little with regard to the interaction of data science experts with functional managers [11, 32, 35]. To the best of our knowledge, the relationship of the two groups, as well as factors shaping the relationship, have not yet been fully examined. This is an issue because companies face several challenges regarding the relationship, e.g. a lack of business objectives for BDA activities, and complex data models that fail business needs [29]. Against this backdrop we pose the following research question: *Which factors jeopardize the relationship between data science experts and functional managers during the collaboration for big data analytics?*

To investigate this relationship, we first conceptualize data science experts and functional business managers as two occupational communities with a different understanding of their work [34]. We argue that a deviation in the respective set of values and norms holds the potential to impede the relationship and thus the important collaboration for BDA. Second, we apply social capital theory to explore the liaison between the two communities and examine the structural, relational, and cognitive relations in order to reveal concrete factors that trouble the relationship. To gain empirical evidence, we conducted a multiple case study where we interviewed representatives from both BDA communities to reflect both perspectives. We point out that a lack of social capital can serve as an explanation for the often problematic relationship between these communities.

2. Conceptual background

As for the conceptual background, we first outline which communities are involved in BDA, how they can be characterized and why the relation might be troubled. Second, we define “collaboration” and illustrate how collaboration in BDA presents itself. Third, we explain social capital theory which we utilize to examine the relationship between the communities.

2.1. Occupational communities in big data analytics projects

We conceptualize data science experts (e.g. data scientists, data engineers) and functional business

managers (e.g. from the marketing or supply chain department) as the two major occupational communities that are involved in BDA projects [15]. In the following, we refer to them as ‘data community’ and ‘business community’ when talking about the individual communities, and ‘BDA communities’ when talking about the sum of them.

An occupational community is a “group of people who consider themselves to be engaged in the same sort of work and who share with one another a set of values, norms, and perspectives” [34, p. 12]. As for the data community, the data scientist who has been acknowledged to have “the sexiest job of the 21st century” [11, p.70] is a well know community member, besides data engineers, and data architects. A data scientist is required to have a very broad skill set that includes a combination of strong analytics expertise and deep business knowledge. Moreover, social competencies are also mentioned as a mandatory skill [32]. The data scientist thus is requested to own a hybrid profile of “a data hacker, analyst, communicator, and trusted adviser” [11, p. 73]. The scarcity of these experts has been identified as a major obstacle in BDA [21]. However, it has been observed that there may be a lack of social skills among data scientists and that they may act solely as ‘number crunchers’ without making any effort to try to build bridges with the business community and share their competencies [32]. This appears as an important weak point within the data community regarding the relationship with functional managers.

The business community in turn can be conceptualized as managers and employees involved in the primary processes of the organization [33], e.g. product managers, sales managers, and supply chain managers. Functional managers are required in BDA projects to utilize data insights in their operations. Specifically, the business community can utilize big data insights for business process improvement, product and service innovation, customer experience enhancement, organization performance improvement, or the creation of symbolic value [14]. When involved in BDA, this community mostly represents the requester and the end user of BDA solutions. As members of this community often have a functional background, they lack a deeper understanding of the analytical methods applied by the data science community, which is i.e. predictive, descriptive, and prescriptive modeling [1]. This appears to be a central weak point within the business community regarding their relationship with the data community. Eventually, these different educational and functional backgrounds outlined above make the overall BDA community, consisting of the data and the business community, a group with a high level of (less visible) diversity [25].

2.2. Collaboration for big data analytics

The word ‘collaboration’ comes from the Latin words *com* (together) and *laborare* (to work). It means that two or more individuals work jointly on an intellectual endeavor [37]. Groups collaborate to create value that cannot be created individually by leveraging diverse skills and backgrounds [6]. Successful collaboration is “the process through which a specific outcome, such as a product or desired performance, is achieved through group effort” [17, p. 40]. In our case, the required collaboration for BDA is threefold: First, data experts need to collaborate within their own community to create excellent data science solutions [13]. Second, business managers need to collaborate within their own community to leverage cross-functional data insights [32], e.g. to gain a holistic view of a customer across all touchpoints. Third, the data and the business community must work together to leverage their technical and managerial skills and achieve business value based on BDA [15]. This third manifestation of collaboration is the focus of this study.

BDA process models and value chain presentations include steps like data collection, preparation, storage, analysis and usage [e.g. 26]. The data preparation and analysis are activities primarily executed by the data community, whereas, in other activities, the business community is also involved (e.g. during the data collection phase by providing data that has been generated in the respective business unit). The last activity, the usage of BDA, means the extent to which BDA solutions are used by the business community to support different business activities [31]. This has been acknowledged as the most crucial antecedent of BDA success [10]. This again underlines the necessity of the involvement of the business community in the BDA process and requires a close collaboration throughout the whole process.

2.3. Social Capital Theory

Social capital is “the sum of the actual and potential resources embedded within, available through, and derived from the network of relationships” [23, p. 243]. It can be understood as the “goodwill that is engendered by the fabric of social relations and that can be mobilized to facilitate action” [2, p. 17]. Social capital constitutes aspects of a social structure and facilitates the actions of actors within this structure [9]. Nahapiet and Ghoshal (1998) distinguish three interrelated dimensions of social capital: the structural, the relational, and the cognitive dimension. First, the *structural dimension* refers to the network of relations as a whole and to the overall connection patterns

between actors (e.g. network ties, identifiable pattern of linkages). Second, the *relational dimension* refers to the quality of relationships between actors that influences the behavior (e.g. trust, expectations, friendship). Third, the *cognitive dimension* refers to shared representations and interpretations, and systems of meaning among groups (e.g. shared language, codes, and narratives). Scholars found that the higher the social capital within a group, the better its performance [4]. This tends to be because the presence of social capital can reduce transaction costs, facilitates collaboration, and enhances mutual commitment [33].

The theory of social capital has been applied to several topics where humans and communities interact, e.g. education, public health, governance, and economic development [16, 27, 39]. In the field of information systems, the theory has been used to examine the relationship between business and IT departments [e.g. 33, 36].

We apply social capital theory as it enables the examination of relationships between communities [5, 33], which is in line with our research objective. We differentiate the social capital according to the abovementioned dimensions in order to gain a deep understanding about the relationship between the data and business community. This relationship appears troubled, and social capital theory can hold an explanation for the obstacles that burden the relationship between the communities and that complicate the collaboration for BDA.

Thus, social capital theory can explain why the performance of the diverse BDA community appears to be rather low, although diversity can also give raise to better group performance [25]. In line with prior studies on the effects of social capital on group performance [5, 23], we suggest that the higher (lower) the level of social capital between the data and business community within collaboration for BDA, the more (less) fruitful the collaboration between these two communities, and thus the more (less) successful the business utilization, and ultimately the higher (lower) the BDA value. The following figure 1 summarizes our conceptual framework for this study.

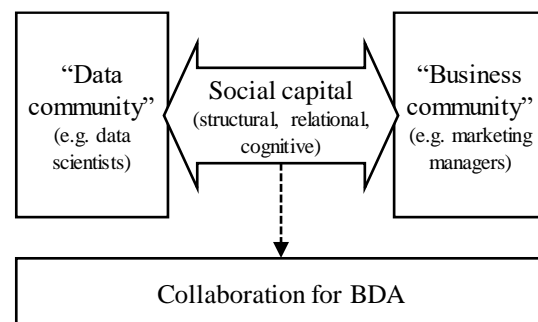


Figure 1: Conceptual framework

3. Research method

As for the method section, we first explain the idea behind selecting case study research for our study, then we explain the rationale behind the case selection, and introduce the case organizations. Lastly, we explain our approach regarding data collection and analysis.

3.1. Case study research

We chose an exploratory multiple-case study research design as this is well suited to examining real life problems in depth in their natural context [40]. Case studies are applicable in the exploratory phase of a topic, e.g. to discover relevant factors of a subject [22]. We want to explore the factors that trouble the relationship between the BDA communities. To ensure a rigorous empirical study, we adhere to the guidelines of qualitative research [12, 24, 40].

3.2. Case selection and sample description

We selected the cases purposefully in a two-step approach. First, we made sure that the selected cases meet important criteria, namely that they apply BDA to derive business value and employ data science experts and functional managers who work with BDA insights. Second, we used a theoretical replication logic to select diverse cases with different collaboration settings (e.g. organizational anchorage of data community in a central team, or distributed within the business units) to allow for contrasting findings [40]. The final selected case sample includes six firms from different industries and with different business models headquartered in Germany. They have all engaged in BDA for at least three years with various business purposes. The representatives from the BDA communities therefore have proven experience in collaboration for BDA. In total, 21 employees (eleven people from the data community, and ten people from the business community) were interviewed by one researcher to avoid bias. Table 1 gives a brief overview of the cases and interviewees, whereas “*” marks the belonging to the data community, and “**” to the business community.

Table 1: Cases and conducted interviews

Case	Attributes (as for 2019)	Interviewees
<i>ManCo</i>	Industry: <i>Manufacturing</i> Revenue: > 80 bn.€	1) manager BDA and machine learning* 2) marketing analytics manager*

	Employees: ~130.000	3) head of quality management** 4) after sales manager**
<i>MediaCo</i>	Industry: <i>Media and Tech</i> Revenue: ~2 bn.€ Employees: ~12.000	1) managing director** 2) head of data science and analytics* 3) senior business consultant** 4) senior manager BI*
<i>RetailCo</i>	Industry: <i>Department store</i> Revenue: ~2 bn.€ Employees: ~32.000	1) area manager IT* 2) area manager BI* 3) department head** 4) procurement analytics manager**
<i>EComCo</i>	Industry: <i>Vehicles</i> Revenue: ~250 mil. € Employees: ~300	1) team lead data science* 2) data scientist* 3) product manager**
<i>ToolCo</i>	Industry: <i>Tools</i> Revenue: > 1bn. € Employees: ~3.000	1) head of data science* 2) data scientist* 3) scrum master**
<i>TrustCo</i>	Industry: <i>Credit</i> Revenue: ~ 200 mil. € Employees: ~900	1) CPO digital products** 2) head of digital product management** 3) head of digital lab*

3.3. Data collection and analysis

Data collection took place from November 2019 to February 2020. We conducted individual interviews which were mostly conducted face-to-face. They lasted from 30 to 70 minutes and were guided by a semi-structured interview guide. All interviews were arranged via the authors’ professional network and were recorded and transcribed verbatim. To build the case database and store, code, and analyze the qualitative data, we utilized ATLAS.ti. Besides performing data triangulation using secondary data (e.g., firm websites, management reports), we had partial access to internal data such as role descriptions and BDA project overviews.

We performed two cycles of coding [20]. First, we assigned descriptive codes as labels to the data to summarize the basic topic of the phrase. This yielded

the jeopardizing factors. Second, we applied pattern coding to group the topics (the jeopardizing factors) to clusters, where social capital served as a theoretical foundation. In unclear cases, fellow researchers joined the discussion. This cycle yielded the assignment of factors to the social capital dimensions. The following table 2 gives an example of the applied coding scheme.

Table 2. Illustration of coding scheme

Quote	Factor (1 st cycle)	Dimension (2 nd cycle)
“One person means something different than he says, at least in the understanding of the other person.”	Different languages	Cognitive social capital
“We are a petitioner for data to work with it.”	Divided data access	Structural social capital
“They start to think of sci-fi movies, it’s a little bit far from reality.”	Unrealistic expectations from business	Relational social capital

4. Results

We present the results of our case study in two steps. First, we outline relevant cornerstones of the BDA collaboration setting within the organizations to allow for a better grasp of the overall situation. Second, we present the identified factors that jeopardize the relationship between the BDA communities. When citing interviewees directly in this section, we use the sorting numbers from table 1 to indicate the source.

4.1. Case descriptions

ManCo is an international manufacturer that is currently undergoing a wide-ranging BDA transformation program. To date, collaboration for BDA is not standardized and has been implemented individually by the departments. During the transformation, the organization aims to install data experts throughout all business functions and data stewards are supposed to act as a liaison role between the communities. The objective of the collaboration for BDA incorporates several use cases, ranging from internal process improvement over databased service development to customer experience improvement.

MediaCo is a tech and media company with several hundred brands. A service provider (100% subsidiary) offers data science know-how to the parent group. The team under the leadership of the head of data science consists of a data scientist, a data engineer, and a

machine learning product owner. This data community collaborates especially with those business units with little or no BDA capabilities. As for the business community, the collaboration is mostly initiated by functional department heads and then executed by functional managers. The objective of the collaboration is multifaceted and ranges from infrastructural projects over reporting dashboards to predictive analysis and natural language processing.

RetailCo is a multichannel department store. The data community is anchored as a central unit and collaborates case-based with the business departments. It consists of two data scientists and two engineers and is led by an IT area manager and a functional area manager. The functional area manager has a bridgehead role with the business units. Within the business community, the sales and procurement teams are the most frequent collaboration counterparts. Business employees in charge for collaborative BDA projects are mostly operational managers. The most common reason for the collaboration is predictive analytics projects, e.g. to predict revenues or to optimize the goods flow.

EComCo is a national online platform for new and used motor vehicles. Its data science community consists of a head of data science who leads a 15-person team, consisting of six data scientists, seven data engineers, a UX designer and a product manager. The team works decentrally in close collaboration with the respective business units, and collaboration tandems are installed across all levels. The collaboration mainly takes place with the advertising, marketing, and sales teams for the provision of big data insights.

ToolCo is an international tool manufacturer that started BDA activities about three years ago. The data science community acts as a central team and consists of a head of data science, a data scientist, a data engineer, and a scrum master. The team to date has collaborated sporadically with the business community across several functions and levels, especially sales. The main goal of the collaboration has been the development of data-based prototypes, e.g. for sales forecasting.

TrustCo is a national credit agency that deals with personal and company data. As (big) data has been at the core of the business model ever since its foundation, the collaboration in the field of (big) data analytics is well established. The data science know-how is bundled within a machine-learning team which collaborates with business units for diverse use cases, but especially with the product management (employees across all levels) in order to develop data-driven products.

4.2. Case analysis

The interviews revealed seven factors that jeopardize the relationship between the data and the business community during collaboration for BDA. These factors can be classified according to the three social capital dimensions. The cognitive dimension of social capital between the communities is characterized by the factors *incongruent mindset* and *different languages*. The structural dimension reveals the jeopardizing factors *lack of joint process*, *divided data access*, *insufficient business alignment*, and *no space for “real” data science*. The factor *unrealistic expectations from the business community* can be assigned to the relational dimension of the social capital between the BDA communities. We will now describe the factors in detail. This is followed by table 3, which gives an overview of the factors (including the frequency of occurrence of the codes in the transcripts, and the number of interviewees that mentioned the respective factor), as well as a reasoning for the theoretical classification of the factors.

Incongruent mindset: The factor named the most often in the interviews and from the highest number of interviewees from both parties is an incongruent mindset between the communities. Put differently, the actors involved in BDA lack a “common data mindset” (EComCo, 1). The incongruence mostly results from the different approaches the two communities have when collaborating for BDA. “Business is often satisfied with ‘a good-enough solution’, whereas data people do not understand this concept, because it has to be right or wrong. And if it's not 100 percent correct, it's wrong” (MediaCo, 3). In order to come closer to a joint mindset, the data community is asked to accept simple data models that can be put into production, instead of those that might yield the highest model accuracy but with no valid business use case.

However, this incongruent mindset is not only prevalent within the data community, but also results from different technological maturities among the business employees. We found that the longer the tenure, the less the trust in technology and the higher the fear of being replaced by machines. “Our tools take a lot of work off the hands, but it is perceived as ‘you’re going to fire me’ (RetailCo, 2). This attitude affects the relationship the business community has with the data community, as these are the people who train the machines and are perceived in a certain way as a cannibalizing factor by the business community.

Different languages. The lack of a shared understanding about the issues which need solving has been put forward both from the data and the business community. The data community on the one hand lacks business understanding, the business community on the

other hand lacks understanding of technology. “They speak different languages, they think differently, and for that reason different results are achieved. One person means something different than he says, at least in the understanding of the other person” (MediaCo, 3). Data scientists demand from business a better understanding “what we need and how we work” (EComCo, 2), but also admit that they are “not aware of all the business projects” (ManCo, 2) in the funnel and consequently cannot always react in a timely manner with the right data. This in turn leads to dissatisfaction in the business community.

Lack of joint process: The lack of a holistic, joint BDA process also hinders a fruitful collaboration. Except for the two cases with a digital business model (EComCo and TrustCo), all case organizations claim that the whole BDA value chain is not specified in terms of handover points and requirements. It is unclear what each community expects from the other. This affects the identification of use cases, the prioritization of those use cases regarding their business impact, the exact definition of requirements, the translation into data science methods and the final execution of the BDA application.

Divided data access: Both communities claim that insufficient data access and ownership complicate the collaboration. Often the data is owned by the business departments which makes it difficult for data teams to do their work. “Someone who only has the magnifying glass will not be able to work with the data if he does not have independent access (MediaCo,1)”. Data science often feels in the position of “a petitioner for data” (RetailCo, 2). Moreover, restricted data ownership within the business community across the business units has also been mentioned as a problem, as it hinders the exchange between functions and thus the generation of cross-department insights. As this can be of special interest for the business community, those data silos prevent collaboration with the data community already at its source.

Insufficient business alignment: Another challenge within collaboration for BDA is the lack of alignment inside the business community regarding data science activities. This factor has two facets: First (put forward by the data community), it relates to the inefficiencies the data community experiences due to reoccurring requests from the business community that have been solved before. “Typical case: I do something for one department, the other department does not know about it. And they are actually trying to solve the same problem in a completely different way. Business people can learn from each other, because in every department there is only one or two guys who can do these things.” (ToolCo,1). The second facet (put forward by the business community) relates to scarce

Table 3: Factors jeopardizing the relationship between BDA communities

Factor (no. of codes / interviewees)	Description	Theoretical reasoning
<i>Cognitive social capital</i>		
<i>Incongruent mindset</i> (11/10)	The business community aims for “good enough” solutions, whereas the data community only accepts 100% correct solutions.	The BDA communities have no shared representation and interpretation of BDA application in business context.
<i>Different languages</i> (11/6)	The business community lacks understanding for technical foundations of BDA, the data community lacks understanding of business reality.	The BDA communities have no shared language, codes and narratives.
<i>Structural social capital</i>		
<i>Lack of joint process</i> (5/5)	There is no standardized process for joint projects. Mutual requirements for the data and business community are unclear.	The BDA communities have no overall connection patterns and lack established patterns of linkages.
<i>Divided data access</i> (4/3)	The data is often owned by the business community, and the data community suffers limitations regarding data access.	The BDA network is not configured properly as data, the most important resource, is not shared.
<i>Insufficient business alignment</i> (4/2)	The data community experiences inefficiencies due to lack of alignment within business community, as it needs to solve similar problems multiple times.	The business community itself lacks established patterns of linkages when collaborating with the data community.
<i>No space for “real” data science</i> (4/2)	The data community cannot excel in BDA results due to limiting restrictions prescribed by the business community within joint projects.	The BDA network is a social structure where the data community cannot prosper.
<i>Relational social capital</i>		
<i>Unrealistic expectations from business community</i> (7/4)	The data community is confronted with unrealistic expectations from the business community regarding possibilities of BDA.	The BDA communities are not relational embedded and cannot leverage mutual bonds and understanding.

resources on the business side. “We all fail because of the same problems. We really can't afford to put so many people in there, somehow you have to create synergies (ManCo, 4)”.

No free space for “real” data science: The data science community complains about having too little freedom to research on big data. Due to time pressure and tight deadlines, complex problems cannot be examined properly and data science activities yield only second-best results. “Even when it fails business needs, that's the nature of research. Doing data research consciously is not meant in the sense of researching in the ivory tower, but with an open end. It is data science, and not data doing” (MediaCo, 3). The business community strengthens this point by opening up another perspective. They criticize the lack of a “data playground” for BDA. “I think the greatest success factor would actually be to have the freedom to use data that is not restricted by regulations. That can happen in a protected room, but that you can basically test everything” (TrustCo, 1).

Unrealistic expectations from business community: The data community claims the prevalence of unrealistic expectations among the business community. Data scientists are confronted with excessive expectations regarding the opportunities of BDA. “Lots of them [business community] go to seminars and advertisements. Most of these statements are not even true. They start to think of sci-fi movies.” (ToolCo, 1). The data community has to demystify the expectation that data science can “simply put relationships within data on the table” (MediaCo, 1) without any input from business. The data community has to educate the business community that “data science is not a miracle machine” (MediaCo, 3). The business community is also blamed for its inaccurate effort estimations. Business either massively overestimates or underestimates efforts for data science activities which yields dissatisfaction in the data community.

5. Discussion

The results of our study provide insights into the relationship between the different occupational communities that are involved in BDA, namely data science and business. The analysis of the jeopardizing factors and the matching with the social capital dimensions show that the relationship between data science and business lacks social capital along all three dimensions, above all cognitive and structural. Based on our findings we suggest the following two propositions:

Proposition 1: The relationship between the data and the business community during BDA projects is troubled due to a lack of social capital. It suffers from different representations and interpretations of BDA application in the business context (cognitive social capital) and the absence of network ties (structural social capital). Specifically, the communities hold incongruent mindsets, speak different languages, lack a joint collaboration process and have no shared data access.

Proposition 2: As the lack of social capital impairs the performance of groups, the troubled relationship is one reason why BDA projects fail. This challenge should be considered in addition to the well-known data, process, and management challenges.

An explanation for the weak social capital shared by the BDA communities can be found by looking at the nature of the lacking social capital compared to similar constellations in the information systems field. Here, the relationship between business departments and classical IT departments can serve as reference. In contrast to the currently emerging BDA communities, the relationship of business and IT departments developed over more than three decades [8]. Van den Hoof and de Winter (2011) found that this established relationship is mostly troubled by a lack of relational capital, followed by a lack of cognitive capital. Structural capital appears not to be an important issue:

As for the *relational capital*, business and IT departments complain about a lack of trust, among other aspects [33]. One could assume that this dimension is not (yet) relevant for the BDA communities, as collaboration has only recently begun and members do not think (yet) about deeper bonds such as friendship etc., where trust would become more important. Assuming that the relationship of the BDA community develops in a similar manner to the business-IT relationship, we propose that once BDA projects have become more common in organizations, it may be increasingly common for the BDA communities to deal with issues related to relational capital as well.

As for the *cognitive capital*, our finding is in line with the finding for the business and IT communities. The business and IT departments mutually feel that the other community does not understand their interests and practices, due to their different skills and mindsets [33]. The incongruent mindset and different languages of the communities result from the different skill sets and personal traits of the community members. As these complementary skills are also needed for the successful execution of BDA projects, this divergence may persist. The “almighty” data scientist who has strong analytics expertise, deep business knowledge and *also* strong social skills, currently appears only as a theoretical solution to this challenge as this combination is hard to find in the job market [21].

As for the *structural social capital*, van den Hoof and de Winter (2011) found that there is a sufficient level of connectedness between business and IT in the organization. This is contradictory to our finding for the data and business community. As the BDA network is rather new, the structural network ties need time to develop to build a strong structural social capital between the BDA communities. One could argue that once BDA projects are more common, the network ties will develop and this lack of social capital will diminish.

6. Implications, limitations and future research

By applying social capital theory to answer our research question “Which factors jeopardize the relationship between data science experts and functional managers during the collaboration for big data analytics?” we contribute to BDA capability and success literature [10, 15, 18, 19, 30] in different ways:

First, we show that the relationship between the two communities data science and business, who need to collaborate during BDA projects, *is troubled*. Herewith, we add to literature that addresses the various skills required during BDA projects [e.g. 11, 15, 21, 32]. Specifically, we contribute that analytical skills of data scientists alone do not bring value unless they are leveraged by the business community, which requires a close collaboration during the BDA process. This collaboration deserves particular attention, as the utilization of BDA insights in business is the most critical contributor to unlock the business value of BDA [10].

Second, and as a further specification of our first finding, we introduce a lack of social capital as possible explanation *why the relationship is troubled*. Specifically, we reveal that the relationship lacks social capital along all three dimension, first of all cognitive and structural social capital, followed by relational

capital. We reveal that there is, above all, a gap in the representation and interpretation of BDA application in business context (cognitive dimension) and the absence of network ties (structural dimension) between the involved BDA communities. Precisely, we discover seven factors that jeopardize the relationship, namely *incongruent mindset* and *different languages (cognitive)*, *lack of joint process*, *divided data access*, *insufficient business alignment*, *no space for “real” data science (structural)*, and *unrealistic expectations from the business community (relational)*.

Third, as we know that the absence of social capital impairs group performance [4], we introduce the troubled relationship between data and business as *one reason why organizations struggle to realize business value from big data*. Thus, we add to literature that is dealing with BDA challenges [30, 35]. We propose to classify the troubled relationship as additional dimension of the management challenges [30]. This category contains, among others, inter-organizational data and information sharing aspects that require “close connections” [30, p. 274], and the troubled relationship between the BDA communities can be seen as intra-organizational challenge in this sense.

Referring to the practical implications, managers should be aware that for the successful implementation of BDA a healthy relationship between the involved communities is essential. As this relationship is still new, managers should *first address the structural dimension* by providing an environment where network ties can develop so that each community member knows whom to reach for a BDA issue and how. A standard approach for BDA projects, virtual task forces, and joint events could be methods of choice. Additionally, managers should take *initiatives to shape a community-overarching BDA mindset*, e.g. by developing a data-driven culture that encompasses aspects such as top management support in formulating data-driven decisions, a data-based operating culture, and the formulation of a company-specific BDA self-perception. Even if the actors of the two communities remain different in terms of mindset and language, such a shared BDA mindset holds the potential to connect the two communities cognitively and potentially merges them to one joint BDA community.

Our study does not come without limitations. A major restriction is that it examined the relation at a single point of time. A longitudinal study that incorporates the dynamics of the social capital would be fruitful. Moreover, the study did not consider external factors that potentially influence the collaboration (e.g. culture, organizational structure, trainings etc.). A detailed investigation would hold the potential to gain a deeper understanding of which circumstances the collaboration holds greater (or even

less) social capital. In order to be even more helpful for practitioners, an examination of the solutions to the identified issues would also be valuable. Lastly, we focused on the data and business community, as they are involved in the primary BDA activities. However, the IT department can also play a role in the collaboration, e.g. when developing an application for a BDA model. Including this community in the discussion might be another interesting avenue for future research.

7. References

- [1] Abbasi, A., Sarker, S., and Chiang, R. H., “Big Data Research in Information Systems: Toward an Inclusive Research Agenda”, *Journal of the Association for Information Systems*, 17(2), 2016, pp. i - xxxii.
- [2] Adler, P. S., and Kwon, S. W., “Social Capital: Prospects for a New Concept”, *Academy of Management Review*, 27(1), 2002, pp. 17-40.
- [3] Asay, M., “Big Data Has Been a Big Disappointment, But There's a Way to Ensure Yours Won't Be”, 2017, retrieved June 22, 2020 from <https://www.techrepublic.com/article/85-of-big-data-projects-fail-but-your-developers-can-help-yours-succeed/>.
- [4] Aquino, K., and Serva, M.A., “Using a Dual Role Assignment to Improve Group Dynamics and Performance: The Effects of Facilitating Social Capital”, *Journal of Management Education*, 29(1), 2005, pp. 17–38.
- [5] Bharadwaj, A. S., “A Resource-Based Perspective on Information Technology Capability and Firm Performance: An Empirical Investigation”, *MIS Quarterly*, 2000, pp.169-196.
- [6] Briggs, R. O., Kolfshoten, G., Vreede, G. J. D., Dean, D. L., “A Seven-Layer Model of Collaboration: Separation of Concerns for Designers of Collaboration Systems”, *ICIS 2009 Proceedings*, 2009, p. 26.
- [7] Chae, H. C., Koh, C. E., and Prybutok, V. R., “Information Technology Capability and Firm Performance: Contradictory Findings and Their Possible Causes”, *MIS Quarterly*, 38(1), 2014, pp. 305-326.
- [8] Chan, Y., und Reich, B., “IT Alignment: What Have We Learned?”, *Journal of Information Technology*, 22(4), 2007, p. 297-315.
- [9] Coleman, J. S., “Foundations of Social Theory”, Belknap Press of Harvard University Press, Cambridge, MA, 1990.
- [10] Côte-Real, N., Ruivo, P., Oliveira, T., and Popović, A., “Unlocking the Drivers of Big Data Analytics Value in Firms”, *Journal of Business Research* (97), 2019, pp. 160-173.
- [11] Davenport, T. H., and Patil, D. J., “Data Scientist: The Sexiest Job of the 21st Century”, *Harvard Business Review*, 90(10), 2012, pp. 70-76.
- [12] Dubé, L., and Paré, G., “Rigor in Information Systems Positivist Case Research: Current Practices, Trends, and

- Recommendations”, *MIS Quarterly*, 27(4), 2003, pp. 597-636.
- [13] Grossman, R., and Siegel, K., “Organizational Models for Big Data and Analytics”, *Journal of Organization Design* 3(1), 2014, pp. 20-25.
- [14] Grover, V., Chiang, R. H., Liang, T.-P., and Zhang, D., “Creating Strategic Business Value from Big Data Analytics: A Research Framework”, *Journal of Management Information Systems*, 35(2), 2018, pp. 388-423.
- [15] Gupta, M., and George, J. F., “Toward the Development of a Big Data Analytics Capability”, *Information & Management*, 53(8), 2016, pp. 1049-1064.
- [16] Jackman, R. W., and Miller, R. A., “Social Capital and Politics”, *Annual Review of Political Science*, 1(1), 1998, pp. 47-73.
- [17] Kotlarsky, J., and Oshri, I., “Social Ties, Knowledge Sharing and Successful Collaboration in Globally Distributed System Development Projects”, *European Journal of Information Systems*, 14(1), 2005, pp. 37-48.
- [18] Mikalef, P., Pappas, I. O., Krogstie, J., Giannakos, M., “Big Data Analytics Capabilities: A Systematic Literature Review and Research Agenda”, *Information Systems and E-Business Management*, 16, 2018, pp. 547-578.
- [19] Mikalef, P., Pappas I.O., Krogstie, J., and Pavlou, P.A., “Big Data and Business Analytics: A Research Agenda for Realizing Business Value”, *Information and Management*, 57(1), 2020, p. 10323.
- [20] Miles, M. B., Huberman, A. M., and Saldana, J., “Qualitative Data Analysis: A Methods Sourcebook” (3rd ed.), Sage Publications, Thousand Oaks, 2014.
- [21] Miller, S., “Collaborative Approaches Needed to Close the Big Data Skills Gap”, *Journal of Organization Design*, 3(1), 2014, pp. 26-30.
- [22] Myers, M. D., “Qualitative Research in Business and Management”, Sage Publications, 2019.
- [23] Nahapiet, J., and Ghoshal, S., “Social Capital, Intellectual Capital, and the Organizational Advantage”, *Academy of Management Review*, 23(2), 1998, pp. 242-266.
- [24] Paré, G., “Investigating Information Systems with Positivist Case Research”, *Communications of the Association for Information Systems*, 13, 2004, pp. 233-264.
- [25] Paul, S., Samarah, I.M., Seetharaman, P., and Mykytyn Jr., P.P., “An Empirical Investigation of Collaborative Conflict Management Style in Group Support System-Based Global Virtual Teams”, *Journal of Management Information Systems*, 21(3), 2004-05, pp. 185-222.
- [26] Phillips-Wren, G. E., Iyer, L. S., Kulkarni, U. R., and Ariyachandra, T., “Business Analytics in the Context of Big Data: A Roadmap for Research”, *Communications of the Association for Information Systems*, (37), 2015, pp. 448 – 472.
- [27] Portes, A., and Sensenbrenner, J. “Embeddedness and Immigration: Notes on the Social Determinants of Economic Action”, *American Journal of Sociology*, 98(6), 1993, pp. 1320-1350.
- [28] Scott, J., “Social Network Analysis: A Handbook”, Sage Publications, London UK, 1991.
- [29] Singh, P., “10 Reasons Why Big Data and Analytics Projects Fail.”, 2017, retrieved January, 2020, from <https://analyticsindiamag.com/10-reasons-big-data-analytics-projects-fail/>.
- [30] Sivarajah, U., Kamal, M. M., Irani, Z., and Weerakkody, V., “Critical Analysis of Big Data Challenges and Analytical Methods”, *Journal of Business Research*, 70, 2017, pp. 263-286.
- [31] Tallon, P. P., and Kraemer, K. L., “Fact or Fiction? A Sensemaking Perspective on the Reality Behind Executives' Perceptions of IT Business Value”, *Journal of Management Information Systems*, 24(1), 2007, pp. 13-54.
- [32] Trolio, G., De Luca, L. M., Guenzi, P., “Linking Data-Rich Environments with Service Innovation in Incumbent Firms: A Conceptual Framework and Research Propositions”, *Journal of Product Innovation Management*, 34(5), pp. 617-639.
- [33] Van den Hooff, B., de Winter, M., “Us and Them: A Social Capital Perspective on The Relationship Between the Business and IT Departments”, *European Journal of Information Systems*, 20, 2011, pp. 255-266.
- [34] Van Maanen, J., and Barley, S.R., “Occupational Communities: Culture and Control in Organizations”, *Research in Organizational Behavior*, 6, 1982, pp. 287-365.
- [35] Vidgen, R., Shawn, S. Grant, D.B., “Management Challenges in Creating Value from Business Analytics”, *European Journal of Operational Research*, 261(2), 2017, pp. 626-639.
- [36] Wagner, H.-T., Beimborn, D., Weitzel, T., “How Social Capital Among Information Technology and Business Units Drives Operational Alignment and IT Business Value”, *Journal of Management Information Systems*, 31(1), pp. 241-271.
- [37] Webster, “Webster’s Dictionary”, Oxford University Press, Oxford, 1992.
- [38] White, A., “Our Top Data and Analytics Predicts for 2019.”, 2019, retrieved December 2019, from https://blogs.gartner.com/andrew_white/2019/01/03/our-top-data-and-analytics-predicts-for-2019/.
- [39] Woolcock, M., “Social Capital and Economic Development: Toward a Theoretical Synthesis and Policy Framework”, *Theory and Society*, 27(2), 1998, pp.151-208.
- [40] Yin, R. K., “Case Study Research: Design and Methods”, SAGE Publications, Thousand Oaks, USA, 2014.